# A GIS-based methodology for identifying pedestrians' crossing patterns

Sylvain Lassarre[a], Emmanuel Bonnet[b], Franck Bodin[c],

Eleonora Papadimitriou[d], George Yannis[e], John Golias[f],

[a] Research Director, French institute of science and technology for transport, development and networks (IFSTTAR), 23 rue Alfred Nobel, F-77447 Marne la vallée, E-mail : sylvain.lassarre@ifsttar.fr

[b] Lecturer, Université de Caen Basse Normandie, UFR de géographie, F-14032 Caen, E-mail : emmanuel.bonnet@unicaen.fr

[c] Lecturer, Université des Sciences et Technologies de Lille, UFR de Géographie et d' Aménagement, Laboratoire Hommes Villes Territoires, F-59655 Villeneuve d' Ascq, E-mail: franck.bodin@univ-lille1.fr

[d] Research Associate, National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Iroon Polytechniou str. GR-15773 Athens, E-mail: nopapadi@central.ntua.gr

[e] Associate Professor, National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Iroon Polytechniou str. GR-15773 Athens, E-mail: geyannis@central.ntua.gr

[f] Professor, National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Iroon Polytechniou str. GR-15773 Athens, E-mail: igolias@central.ntua.gr

**Abstract**

A pedestrian trip is a spatiotemporal process going through different states and related to different decisions made at certain times and locations on the urban network. The analysis of pedestrian trips in terms of crossing patterns is a complex task, which is often further limited by a lack of appropriate and detailed data. The objective of this research is the development and testing of appropriate indicators of pedestrian crossing behaviour along urban trips, and a methodology for collecting and processing the data required for the analysis of this behaviour. First, a comprehensive set of indicators for the assessment of pedestrian behaviour in urban areas is proposed (i.e. average trip length, number, type and location of crossings). Then, a GIS tool is developed for the storage and integration of information on pedestrian trips, and the crossings made during the trips, with other geographical information (e.g. road network function and geometry, traffic control and pedestrian facilities). The proposed approach is then tested at network level on a sample of pedestrian trips collected by a field survey. The results suggest specific patterns of pedestrian crossing behaviour, such as the tendency to cross at the beginning of the trip and the tendency to cross at mid-block locations when signalized junctions are not available. The results are further discussed in terms of urban planning and management implications. It is concluded that the proposed approach is very efficient for the analysis of pedestrian crossing behaviour in urban areas.

**Key-words**: pedestrian trips; pedestrian behaviour; geographic information; data integration.

## 1. Introduction

The analysis of pedestrian mobility and behaviour in urban areas may enable the testing of alternative engineering and traffic control strategies, and the estimation of the level of service and road safety of pedestrians. In particular, having information on the number, type and location of crossings made during urban pedestrian trips may be very useful for the planning, implementation and evaluation of interventions aiming to create an uninterrupted and self-explaining walking environment, with improved design and operational elements, such as protected crossings and other pedestrian facilities (Papadimitriou et al. 2010).

Measuring pedestrian activity in urban areas is a complex task. The analyses of pedestrian movements in urban areas are mainly devoted to route choice (Kukla et al. 2001; Hoogendoorn & Bovy, 2004; Antonini et al. 2006), activity scheduling models (Borgers & Timmermans, 1986; Kitazawa and Batty, 2004), pedestrian flow or crowd models (Lee and Lam 2008; Daamen et al. 2005) and evacuation models (Xiaoping et al. 2009).

A lot of research is devoted to behavioural aspects of pedestrian activity; however, different indicators of pedestrian behaviour are used in different studies, whereas a framework for analysing pedestrian behaviour in urban areas in a complete and systematic needs to be considered (Papadimitriou et al. 2009). The behavioural indicators used are most often macroscopic, such the number of trips (Jonah & Engel, 1983) or the number of crossings (Howarth, 1982; Keall, 1995), the time spent walking (Lee & Abdel-Aty, 2005) and the distance travelled (Baltes, 1998). Moreover, several indicators dedicated to the crossing behaviour of pedestrians in urban areas have been proposed, including the selection of crossing location, such as the probability of crossing at junction or at mid-block (Chu et al. 2003, Baltes and Chu, 2002), the acceptance of traffic gaps , such as the distribution of traffic gaps or delays accepted (Hamed, 2001, Das et al. 2005), the compliance to traffic rules, such as the proportion of crossings outside the crosswalk area or during the red signal (Sisiopiku and Akin, 2003, Yang et al. 2006). However, most of these indicators concern crossing behaviour at local level, whereas crossing patterns along entire trips are seldom explored (Lassarre et al. 2007).

Moreover, the number of trips or the time spent during a walking trip is not precise enough to assess the pedestrian's exposure to the risk. In fact, a pedestrian is mainly exposed to the risk of road accident when crossing a road (Duncan et al. 2002), therefore the detailed recollection of crossings made by a pedestrian and the characteristics of the traffic flows crossed is the basis to assess pedestrians' exposure (Lassarre et al., 2007).

Within the various methodological issues that need to be addressed for analysing pedestrian crossing behaviour during urban trips, one of the main limitations is the lack of data and the relative difficulty in obtaining detailed and reliable data along real pedestrian trips. Most researchers use data from video recordings, which can be very efficient for a detailed description of local behaviour (Bierlaire et al. 2003); however, they can not be used for network-level analysis purposes. Other researchers use interview results (Hine 1996, Fitzpatrick et al. 2004), which can yield attitude rather than behavioural data and often suffer from the known limitations of self-reporting. In recent research, field survey techniques are proposed, which are based on following pedestrians and recording their behaviour by means of a camera in motion (Papadimitriou et al. 2010); despite the advantages of such techniques, significant effort is required for processing and coding the important amount of information collected.

Within this context, Geograghic Information Systems (GIS) can be used to code, visualize and process spatial information in a very detailed and accurate way and could be exploited for the analysis of pedestrian behaviour along urban trips (Li and Tsukaguchi, 2005). Most existing pedestrian-related GIS applications concern either macroscopic evaluations of pedestrian walking space, walkability and accessibility in urban areas (Aultman-Hall et al. 1997; Crane & Crepeau, 1998; Schlossberg & Brown 2004; Parks & Schofer, 2006) or mapping and identification of high accident risk sites for pedestrians (Miller 2000; Bajleri et al. 2007; Pulugurtha et al. 2007).

The use of GIS for the analysis of pedestrian crossing behaviour in urban areas could provide an advantageous combination of sufficient level of detail, accuracy and flexibility. The integration and spatial matching of quantitative and geographical field survey data on pedestrian behaviour with road and traffic information in a GIS, would allow for estimation of various pedestrian behaviour indicators in a straightforward way, as well as for detailed analysis of this behaviour in terms of the effects of the road and traffic environment. Moreover, the possibility of visualising the analysed information offers additional advantages for the exploitation of the results in urban planning and management.

According to the above, this research has a twofold objective: first, to propose a set of appropriate indicators for the assessment of pedestrian crossing behaviour in urban trips and second, to develop a GIS-based survey framework for the collection and processing of the necessary data at trip level. Both contributions are then tested by means of a field survey in Lille, France, allowing to identify distinct pedestrian crossing patterns in the study area and to demonstrate the usefulness of the proposed approach.

## 2. Methodological contributions

### 2.1. Selection of pedestrian behaviour indicators

Pedestrian trips may be more or less complex and include several changes of direction. A first problem, therefore, is to transform a complex trip from an origin to a destination into an appropriate form, which could be analyzed in a more straightforward way. Another is to localize, count and discriminate the different types of crossing options (crossing a road at a junction or at mid-block, on or out of a pedestrian crosswalk etc.).

In the present research, it is shown that trips with several changes of direction can be handled similarly to trips with no change of direction, because the properties of road crossings are largely defined by whether they take place across or along the 'trip roadway', regardless of the number of changes of direction. The trip roadway is here defined as the sequence of road links of the road network, along which the pedestrian trip takes place. Accordingly, the path of a pedestrian trip can be either parallel to the trip roadway (i.e. the pedestrian is walking along the roadway on sidewalks) or intersect with the trip roadway (i.e. the pedestrian is crossing the roadway).

On the basis of a topological analysis approach, the roadway of a pedestrian trip may be considered to divide the plane in two distinct parts, an 'interior' and an 'exterior', as shown in Figure 1. The origin and the destination of the trip are thus situated on either part of the plane. It can be then shown that, in trips with origin-destination on the same part of the plane (labelled O/D=1 from this point on) the pedestrian path will intersect the trip roadway an even number of times, while trips with origin-destination on different parts of the plane (labelled O/D=2 from this point on) will intersect the trip roadway an even number of times[1].

***Figure 1 to be inserted here***

Based on the above, two categories of crossings can be considered, taking into account the road network layout and the pedestrian trip configuration (Lassarre et al. 2007; Papadimitriou et al. 2010):

- "Primary" crossings made across the trip roadway, at junctions or mid-block locations.
- "Secondary" crossings made across roads perpendicular to the trip roadway, while moving along the trip roadway, at junctions.

[1] This intuitive consideration is based on the 'Jordan curve theorem' of topology; for details the reader is referred to Papadimitriou et al. (2010).

It has been demonstrated that the number of primary crossings is probabilistic, in the sense that a pedestrian may choose between several options (junctions, mid-block crosswalks, other mid-block locations etc.) available along the roadway, while the number of secondary crossings is deterministic, in the sense that each secondary crossing will take place at a given junction, either in the 'interior' or the 'exterior' part of the plane, as a consequence of following the particular path (Lassarre et al. 2007). For example, in Figure 1, in path (a) a primary crossing takes place across the second road segment of the trip roadway; therefore a first secondary crossing takes place in the exterior part of the plane. However, in a path (b) for the same trip roadway, a primary crossing takes place across the first road segment, so a first secondary crossing takes place at the same junction, but in the 'interior' part of the plane. In any case, a second secondary crossing takes place in the interior part of the plane just before the destination.

As mentioned above, the total number of primary crossings along a trip depends on the origin-destination combination in relation to the trip roadway. In particular, if the origin and destination are on the same part of the plane (interior or exterior, as defined above), a pedestrian would not have to cross; if the origin and destination are on different parts of the plane, a pedestrian would have to cross at least once in order to reach his destination. In complex trips with several changes of direction, pedestrians are likely to make additional primary crossings, in order to minimize walking distance. Therefore, the total number of crossings N along a pedestrian trip may be described by a probability distribution, in relation to the O/D configuration of the trip (Lassarre et al. 2007).

As regards the location of crossings, two levels can be identified: the first one concerns the distribution of crossings along the trip. For instance, recent research results suggest that pedestrians tend to make the necessary primary crossings to reach their destination at the beginning of the trip, rather than towards the end of the trip (Lassarre et al. 2007). It has also been found that pedestrians prefer to "cross first and walk later" (Chu et al. 2003). Consequently, the distribution of the number of crossings in relation to the distance from the trip origin would be a useful indicator of pedestrians crossing behaviour, as it could reveal specific areas with increased expected number of crossings e.g. around a pedestrian trip generator.

Furthermore, each crossing decision involves a choice between a set of specific options offered to the pedestrian. In particular, in each road segment, two junction options are available (one junction at the beginning and one junction at the end of the road segment) and one mid-block location defined as the area between the two junctions. Within this framework, it is noted that Junction 2 of a road segment practically coincides with Junction 1 of the next road segment (see Figure 1). At the moment of the crossing decision, the pedestrian shall have to identify the options available nearby (junctions - traffic controlled or not, mid-block locations - protected or not) and select one option (Chu et al. 2003, Lassarre et al. 2007). Therefore, the comparison between the options available around each crossing with the characteristics of the location of the crossing, may provide additional insight in the way that pedestrians make crossing decisions.

According to the above, a set of indicators for the identification of pedestrians crossing patterns can be considered as follows:

- Number of primary crossings per trip
  - If origin-destination are on the same side of the roadway, then N {0, 2, 4, …}
  - If origin-destination are on different sides of the roadway, then N {1, 3, 5, …}
- Number of secondary crossing per trip
- Distribution of primary crossings in relation to the distance from origin
- Distribution of the type of crossing options available
  - Junctions or Mid-block
  - Protected or not
- Distribution of the type of crossing locations selected
  - Junctions or Mid-block
  - Protected or not

These indicators, as well as the related probability distributions, can be estimated on the basis of the analysis of a sample of collected trips. In the next sections, it will be shown that this type of analysis can be based on a database of the road network and a GIS methodology for matching the trajectories of pedestrians on the digitized road network.

## 2.2. Exploitation of GIS

### 2.2.1. Data collection methods for pedestrian trips

In order to collect detailed data on pedestrian trips, two techniques can be used: either by equipping pedestrians with a GPS device and recording automatically the trajectory of the trip (within the limits of precision and the failures of the device), or by following pedestrians on the road, with or without their knowledge, and recording the trip by drawing the trajectory on a map. In both cases, the data can be coded and processed into a spatiotemporal referential geographical system.

When using a GPS device, one can obtain (x,y) coordinates at regular time steps. The trajectory is then automatically recorded, by joining the locations collected in time, and geo-referenced into a GIS. At the semantic level, for each point of the trajectory, one gets time related information such as the time elapsed since the beginning of the trip, an exact position, and an instantaneous speed. From that aspect, GPS data are faster to obtain and to process.

However, when drawing the trajectories manually on a map, one can obtain a continuous line following the "exact" trajectory (plus the times of crossings if a chronometer is used). Moreover, accuracy problems may be encountered with GPS devices, which are due to the variation in the reception of the signals in dense urban areas. During a pilot survey carried out within the present research, two identical GPS devices were used in parallel for the same trajectory (i.e. a person was equipped with two devices during a 5 minute walk), and some non negligible deviations were observed. These deviations mainly concerned the accuracy in the (x,y) coordinates recorded, which may under certain conditions affect the number and location of crossings recorded (e.g. misclassify mid-block crossings as junction crossings).

Although solutions to minimize these distortions are available nowadays, it was quite difficult to obtain the amount and level of detail and accuracy of the necessary information for analyzing pedestrians crossing patterns by means of GPS at the time the survey was carried out (i.e. in 2004).

Consequently, within the present research, the following of pedestrians was opted for, by drawing their trajectory on an orthomap and by recording on a sheet the information related to the trip. In the following sections, it is shown how this method is efficient for an in-depth recording of pedestrians' behaviour in urban areas.

### 2.2.2. A database for pedestrian urban trips

The data on pedestrian behaviour in the urban road network needs to be appropriately structured and processed, so that it can be stored in a GIS. Part of the data comes from official data files concerning the road network; however a lot of complementary information needs to be collected, especially as regards pedestrian crossings.

In France, geographical data come from the National Geographical Institute (IGN), which provides objective and official information on the territory. In order to analyse the behaviour of pedestrian, it is important to have complete information on the road network and the traffic conditions (e.g. traffic volumes, traffic control). Moreover, new geographical data, not provided by IGN, has to be created and coded (e.g. pedestrian crossing facilities, such as crosswalks).

Therefore, the GIS designed in the present research is composed by 5 information layers:

- A layer describing the accessibility (function) of the road network, in which road segments are classified into three categories: pedestrian only, vehicle/pedestrian, vehicle only (see Figure 3). Moreover, a standard graph representation (arcs and nodes) is used. When a road segment is vehicle/pedestrian, it is further specified whether both sides of the road are accessible for pedestrians, or whether only one side (left or right).
- A layer describing the geometry of the road network, in which the following attributes are recorded for each road segment: number of directions (i.e. one-way or two-way road), road width, number of lanes in each direction, presence of lane separation.
- A layer concerning the traffic lights. These are coded as additional nodes located on the road network and their attributes include the type of location (junction or mid-block), and the traffic streams covered (e.g. separate left-turn).
- A layer including the marked crosswalks, also coded as additional nodes on the arcs representing the road network. Each crosswalk is defined by a set of attributes including the number of lanes to cross, the nature of the protection offered (median) and the type of the crosswalk (zebra, pelican, bridge, tunnel). This information is obtained from aerial photographs.
- A layer including information on sidewalks (shoulders) for each road.

## 3. Data collection and processing

### 3.1. Collection of data on pedestrian trips

The survey site was the area around the metro station "hotel de ville" (town hall) in the Villeneuve d'Ascq suburban area of Lille, France. The study area is a new city and focus is put on its southern part, which is the most active and populated one, and includes three metro stations and a University campus. A map of the study area is presented in Figure 2.

***Figure 2 to be inserted here***

For the purposes of the analysis, 83 adults (20-65 years old) were followed at the end of March 2004 on weekdays at different time of the day (morning, evening, off peak hours). Each individual was followed from the exit of the metro station without his or her knowledge, during 5 minutes maximum, given that pedestrian behaviour in longer trips may not be typical. There were a little bit more women than men in the sample (54 %) and most of the pedestrians were alone on the trip.

In order to code the trajectory information, orthophotographs were used. These images have a great spatial resolution (a pixel by 0.16 m), providing a very precise localisation of the individual, both on sidewalks and crosswalks. For each following, one surveyor noted the trajectory and another surveyor collected the information related to the crossings made by the pedestrian.

### 3.2. Collection of geographical data

The urban environment in the study area is dedicated to administrative tertiary sector, trade, services of proximity, culture (forum of sciences) and residence, and also includes a School of Architecture and a primary school. As regards the accessibility of the road network, all three functions were identified in the study area, as shown in Figure 3.

***Figure 3 to be inserted here***

The GIS tool developed followed the five-layer structure outlined in section 2.2.2. The layers concerning the road network accessibility and geometry, as well as the traffic control information, were obtained by the official IGN data. The main source of additional information for the remaining layers was the use of orthophotographs and the field survey data; it is noted that the most interesting aspects of pedestrian behaviour were observable either on the sidewalks or at the crossing locations. The details of the GIS tool are presented in Table 1.

***Table 1 to be inserted here***

### 3.3. Integration of the trajectories and the crossings of pedestrians

Two operations need to be undertaken once the data are introduced into the GIS: the transformation of the raw trajectories into polylines (i.e. a digitalization of the trips traced on the orthophoto during the field observations), and the spatial matching for the identification of the crossings and their characteristics.

### 3.3.1. Transformation of the raw trajectories into polylines

From the survey, 83 pedestrian trips were drawn on the orthophoto views and reported in a GIS under the form of a georeferenced object. Each trip was inserted as an additional separate layer (totalling a set of 83 layers), which could be combined with the other layers of the GIS tool (road network, crosswalks etc.), and examined in relation to the other geographic objects of the road network. The trips are georeferenced by positioning them in accordance to the field characteristics on the basis of matching (x,y) coordinates of the orthophoto with the GIS. Due to the great precision of the orthophoto, sidewalks are easily identified and one can directly track the trips on them. The trajectory is represented as a polyline (Figure 4), from which the distance of the trip and the average walking speed can be calculated. It is noted that a few trajectories are truncated at the end, which is the case when the trip lasts more than 5 minutes (i.e. the maximum duration allowed for following).

***Figure 4 to be inserted here***

### 3.3.2. Spatial matching

The next step for the integration of the survey information consists in identifying and characterising the crossings made by each individual, and by creating an additional node for each crossing. All the crossings are obtained by just one operation: intersecting two polylines (i.e. the pedestrian trip polyline and the road network polyline). At each intersecting point identified, a point (node) is created and added to the pedestrian trip polyline (Figure 4). This can be implemented either manually or automatically, by means of appropriate commands. A Table with the attributes of each crossing is also automatically attached to the geographical coordinates of its position, including some geographical information of the road crossed (i.e. name of the road, width of the road, traffic volume etc.) with a spatial join.

In addition to the information created by means of the automatic spatial matching, additional information on each crossing can be coded, either automatically or manually; this concerns the indicators of pedestrian behaviour (primary/secondary crossing, junction or mid-block, distance from origin etc).

Summarizing, the spatial matching operations follow the following algorithm:

A) Identification of the trajectory as a polyline (automatically)

   From the trip origin, connect to the network,

   From the trip destination, connect to the network,

   Do from origin to destination

   Find the road segment, the direction and the side.

B) Identification and characterization of the crossings (automatically or manually)

   Identify the points of crossing (a cell or a pair (x,y)) from the intersections of trajectories with the network

   Look for the location: junction/mid-block

   Check if there is a traffic light

   Check if there is a crosswalk

   Characterize the crossing as primary or secondary

It is noted that, in the present research, the second operation was done manually, due to the limited number of crossings to be processed. Nevertheless, it could be also done automatically in a rather straightforward way, by implementing the appropriate commands, allowing to identify the main trip roadway and classify crossings into primary and secondary, as described in section 2.1.

## 4. Implementation and results

On the basis of the geographical data collected, a typology of pedestrian crossing behaviour can be created based on the proposed set of indicators. First, basic characteristics of pedestrian trips are analysed (i.e. length, duration, walking speeds). Moreover, the number and type of crossings per trip are examined. Finally, a breakdown of crossing decisions per junction or mid-block, protected or non-protected is created.

### 4.1. Length and duration of pedestrian trips

The minimum length of the pedestrian trips monitored in the study area was 66 metres and the maximum length was 800 metres, whereas the average trip length was 350 metres. By considering an average walking speed for adult pedestrians equal to 1.35 m/sec, an average walking time equal to 4.3 minutes can be estimated (with minimum 0.8 minutes and maximum 9.8 minutes). It is noted that the measure of the trip's speed is composite because it includes both the walking and the waiting time, as in Equation (1).

Average trip speed = (walking time along the sidewalk + walking time crossing the road + waiting time before crossing + other waiting time on the sidewalk) / (length of trip along the sidewalk + length of the crossings) (1)

The speed is therefore expected to depend on the age, the gender of the pedestrian and on whether the pedestrian is walking alone or in a group.

By superimposing an appropriate grid over the study area (e.g. 3 × 3 metres cells, with 3 m the average width of a sidewalk), and by counting the number of passages through each cell, it is possible to calculate the related densities and visualize the overall distribution of the sampled pedestrian trips from the metro station (Figure 5) in the study area. An advantage of this technique is that am immaterial and permanent in time and space geographical layer is obtained, which allows for precise and detailed analysis of the territory and the processes.

***Figure 5 to be inserted here***

From the distribution of the sampled pedestrian trips around the metro station, it can be noticed that pedestrian density does not decrease with the distance from the metro station. More specifically, pedestrian trips are not clustered around the metro station, as might be expected, but are in their majority extended along specific axes of the examined road network. This finding suggests that infrastructure design and traffic control (e.g. in terms of sidewalk width, marked crosswalks and pedestrian signals) should be implemented beyond the immediate proximity of the metro station.

Moreover, it can be seen that increased pedestrian density is observed along roads with decreased crosswalk density, suggesting that the placement of marked crosswalks in the study area needs to be improved in order to meet the actual needs.

**4.2. Number and type of crossings**

The frequencies of crossings per trip are summarized in Table 2. The survey results confirmed the initial assumption that, in trips where origin and destination are on the same part of the plane (interior or exterior, as defined on the basis of the main trip roadway), the number of primary crossings is even {0, 2, 4, …}, whereas where origin and destination are on different parts of the plane, the number of primary crossings is odd {1, 3, 5, …}. Moreover, pedestrians appear to minimize the total number of primary crossings per trip. It is also interesting to note that all trips include at least one secondary crossing.

***Table 2 to be inserted here***

In order to model the expected number of crossings during a pedestrian trip, various probability distributions for discrete data were fitted to the frequencies of primary and secondary crossings per trip (including Binomial, Poisson, Negative Binomial, Integer Uniform and Geometric distributions). It is noted that the knowledge of the probability distribution of the number of crossings per pedestrian trip may be combined with the information on trip density in the examined area, for more efficient and targeted planning of pedestrian facilities.

Given the particularity of the primary crossings consideration, an appropriate theoretical approach was opted for. First, the sample values (primary crossings per trip) were truncated to 4, excluding the single case of 5 primary crossings per trip. Then, given that the values {0, 2} refer to trips with side O/D=1 and the values {1, 3} refer to trips with side O/D=2, the two sub-cases were examined separately. In this framework, it is quite straightforward to consider that both sub-cases correspond to Bernoulli trials with two possible outcomes, and thus follow Binomial distributions.

More specifically, the number of primary crossings in trips with side O/D=1 corresponds to a Binomial distribution with parameters (n=45, p=0.489), where n equals the total number of trials and p equals the expected probability of successful outcome (considering that the case of 2 primary crossings is the "successful" outcome against 0 primary crossings). Accordingly, the number of primary crossings in trips with side O/D=2 corresponds to a Binomial distribution with parameters (n=38, p=0.132), where n equals the total number of trials and p equals the probability of successful outcome (considering that the case of 3 primary crossings is the "successful" outcome against 1 primary crossing).

Therefore, the distribution of the total number of primary crossings is a combination of two Binomial distributions, which corresponds to a multinomial distribution with parameters (n=82, $p_1$=0.280, $p_2$=0.390, $p_3$=0.269, $p_4$=0.061), where n is the total number of trials and {$p_1$,…, $p_4$} are the expected probabilities for each possible outcome (Figure 6).

***Figure 6 to be inserted here***

As far as the frequency of secondary crossings per trip is concerned, this is independent of the O/D type and can be intuitively considered to follow a Poisson distribution with parameter $\lambda$=2.94, which corresponds to a chi-square goodness-of-fit statistics equal to 1.58 with a p-value equal to 0.812. However, given that the zero value is not available in the dataset (i.e. all trips include at least one secondary crossing), a more appropriate distribution for this dataset may be a left truncated Poisson distribution (Cameron and Triverdi, 1998, Bancroft et al. 1983), which is defined as in Equation (2):

$$P(X=x) = \begin{cases} 0 & \text{if } 0 \leq x \leq m-1 \\ \dfrac{\lambda^x}{x!} \dfrac{1}{\sum_{x=m}^{n} \dfrac{\lambda^x}{x!}} \quad (x = m, \ldots, n) & \text{if } m \leq x \leq n \end{cases} \qquad (2)$$

In the case when just the zero value is excluded, as is the present case, the mean of the left truncated Poisson distribution is:

$$E(X \mid X > 0) = \frac{\lambda}{(1 - e^{-\lambda})}$$

As the mean of the observed distribution of the number of secondary crossing is equal to 2.94, the mean of the left truncated Poisson distribution is estimated from the above equation equal to $\lambda$ = 2.76. Testing the observed frequencies against this distribution gives a chi-square goodness-of-fit value equal to 0.890 with a p-value equal to 0.827.

It is thereby indicated that a left truncated Poisson distribution provides an improved fit for the number of secondary crossings per trip (Figure 7).

***Figure 7 to be inserted here***

### 4.3. Distribution of crossings within the trip

Another indicator of pedestrian behaviour is the location of crossings within the trip i.e. the tendency of individuals to cross earlier or later along the trip. This analysis only concerns primary crossings, as the secondary crossings are only a consequence of the choice of primary crossings. The results of the survey are presented in Figure 8.

***Figure 8 to be inserted here***

It is interesting to note that 25% of primary crossings take place during the first 20% of the trip length and 50% of primary crossings take place during the first 40% of the trip length, regardless of the relative setting of origin and destination. This strongly suggests that pedestrians are more likely to cross earlier along the trip, i.e. reach the side of the roadway corresponding to their destination as soon as possible and then carry out only the necessary secondary crossings. In particular, in trips with origin and destination on the same side of the plane (O/D=1), where it was shown that pedestrians are more likely to carry out two crossings, the crossings are distributed along the trip in a more uniform way. More specifically, there is a higher probability of crossing towards the end of the trip compared to the average, obviously due to the fact that there is a second crossing to be carried out. On the contrary, in trips with origin and destination on different sides of the plane (O/D=2), where it was shown that pedestrians are more likely to carry out one crossing, 75% of primary crossings take place on the first 60% of the trip length, while the probability of crossing towards the end of the trip is lower compared to the average.

The above probability distributions can be exploited for weighting the number of crossings along the related trips in relation to the total trip length. For example, it is deduced that the distribution of primary crossings is not expected to be uniform along the road axes of increased trip density identified in Figure 5, and therefore the priority in the implementation of crossing facilities needs to be given within an estimated range from the metro station in which primary crossings are more likely to occur.

### 4.4. Distribution of crossings between mid-block and junction

The selected location of a primary crossing is an important determinant of pedestrian behaviour. The decision between junction and mid-block location, in combination with the related selection of a protected or unprotected one, may vary significantly between different individuals. In Figure 9, the decisions of pedestrians in the study area are examined in relation to the respective options available.

***Figure 9 to be inserted here***

In particular, on each road link where a primary crossing took place, there were three options available, junction 1 (i.e. the first junction on the pedestrian's way), mid-block and junction 2 (i.e. the second junction on the pedestrian's way). It is reminded that mid-block corresponds to the area between the two junctions, and consequently nearly infinite crossing location options may be considered in theory; in practice, the differences in road and traffic conditions between two consecutive junctions will be in most cases negligible, allowing to consider a single mid-block option for the whole area between the junctions.

The right panel of Figure 9 shows the proportion of protected options for junction or mid-block. A protected option refers to the presence of a marked crosswalk or a traffic signal. The left panel of Figure 9 presents the breakdown of crossing decisions of pedestrians per crossing option.

Only 15% of all crossing decisions concern junction 2; the rest of the crossing decisions are evenly distributed between junction 1 and mid-block. More than one third of all the options available in the study area were protected ones, and this proportion is similar for all types of options (junction and mid-block). Surprisingly, the results show that pedestrians prefer non-protected over protected crossing options in general. In particular, the distributions of non-protected vs. protected crossing decisions are in accordance to the respective distribution of the options available only for junction 1; specifically, the odds ratio of non-protected decisions to non-protected options is equal to 0.95 in this case.

On the contrary, at mid-block and at junction 2, non-protected crossing decisions are clearly over-represented in relation to the related options available (the odds ratios are equal to 3.62 and 2.85 respectively). With respect to mid-block options, this may be attributed to the fact that pedestrians feel less protected at mid-blocks in general, or they sense that vehicles are less likely to yield for them at mid-block, regardless of the presence of pedestrian crosswalk, and consequently they do not distinguish between protected and non-protected mid-block locations. It is also noted, that very few traffic-signal-protected options were available at mid-block in the study area. As regards junction 2, this may be explained as follows: given that junction 2 of a road link is strongly affected by junction 1 of the next road link, it is possible that pedestrians crossing at junction 2 do not distinguish between these two options and feel protected by the facilities of junction 1.

In general, from the above results, it can be deduced that pedestrians are more likely to cross on the first option along their way (junction 1 or mid-block), however they are less cautious at mid-block locations and appear to opt for non-protected ones.

## 5. Discussion

The results of this research concern an in-depth empirical analysis of pedestrian crossing behaviour along urban trips. The proposed indicators and the related probability distributions reveal specific crossing patterns over a sample of pedestrians with a given activity at network level, as regards the length and density of trips in the study area, the number of crossings (by type) per trip, the location of these crossings within the trips, and the type of location for each crossing decision. These indicators may be used for testing alternative engineering or traffic control strategies as per the behaviour and safety of pedestrians in the study area. For example, the analysis of the proportion of pedestrians crossing at mid-block reveals the need for improving the design and operational elements at junctions or for creating more mid-block crossing facilities.

Moreover, the knowledge of the distribution of pedestrian crossings in relation to the distance from the trip origin e.g. the metro station, reveals the need for area-wide interventions for the improvement of pedestrians level-of-service and safety, beyond the immediate proximity from the metro station. An important indicator was also defined and estimated: the distribution of the number of crossings made along a trip according to their type: primary or secondary. Given that several secondary crossings are necessary for following a particular path (even a path of zero primary crossings), facilities need to be implemented in a way to address the need for crossing movements both across and along pedestrian trip roadways.

Moreover, the case-study of Lille indicated specific road crossing practices, such as the relationship between the number of crossings (odd or even) and the origin-destination configuration on the road network, the tendency to cross roads at the beginning or the trip rather than later on, and the unattractiveness of mid-block marked crosswalks for pedestrians. These results certainly reflect particularities of the specific network (few traffic signals, low traffic volumes etc.). They also reflect the socio-demographic composition of the pedestrians' sample e.g. women tending to be more cautious when crossing than men, or senior people choosing protected crossings as much as possible, which were not given particular emphasis in the present research.

It would therefore be interesting to investigate these issues on a more typically urban network, with higher traffic volumes and more diverse land use. It is underlined that traffic volume data can be integrated in the proposed GIS tool as an additional layer. It would then be possible to examine the variation in the proposed pedestrian behaviour indicators in relation to traffic conditions. Furthermore, it would be interesting to examine pedestrian crossing behaviour in relation to trip purpose, as well as in relation to whether it is a pedestrian-only trip or a complementary pedestrian trip with other transport modes (e.g. car, public transport etc.).

The use of a GIS was proved to be very efficient for the localization and coding of the information on trips and road crossings of the survey sample. The combined use of orthophotographs and official geographic data allowed for the development of an ad hoc five-layer tool. Moreover, the method of following pedestrians, rather than using video recordings or GPS device outputs, allowed not only for the collection of very detailed and accurate information but also for the storage and processing of the information in a quite straightforward way. Although the whole process can be implemented automatically, by means of additional computer programming, part of the spatial matching of pedestrian trips and geographic information can also be carried out manually, which may be proved to be more cost-effective with small samples as the one in the present research.

The methods proposed in this research seem promising for the identification of pedestrian crossing patterns in urban areas. A new survey method based on the exploitation of GIS properties in pedestrian behaviour research was developed and tested, focusing on the aspect of road crossing and addressing several existing difficulties involved in obtaining detailed and accurate data on pedestrian trips. Moreover, a comprehensive set of pedestrian behavioural indicators was proposed, covering all important aspects of pedestrian behaviour, including indicators used in previous studies as well as new ones, and meeting the need for analysis at trip level.

The proposed set of indicators could certainly be extended to include other information e.g. the degree of compliance to the traffic rules and yielding behaviour. Such indicators could be easily added to the set of indicators on the crossing choices of pedestrians (junction or mid-block).

The results can be exploited for urban planning and interventions to improve pedestrian level of service e.g. assessment and comparison of different urban planning scenarios with respect to the implementation of pedestrian crossing facilities at an area-wide level.

**Acknowledgement**

This work has been supported by the European Commission within the "HEARTS - Health Effects and Risks of Transport Systems" Project of the 5th Framework Programme.

**References**

Aultman-Hall L., Roorda M., Baetz B.W., 1997. Using GIS for Evaluation of Neighborhood Pedestrian Accessibility. Journal of Urban Planning and Development 123 (1), pp. 10-17.

Antonini G., Bierlaire M., Weber, M. (2006). Discrete choice models of pedestrian walking behavior. Transportation Research Part B 40, pp. 667-687.

Baltes, M., Chu, X., 2002. Pedestrian level of service for mid-block street crossings. In: TRB 81st Annual Meeting, Transportation Research Board, Washington DC.

Baltes, M.R. (1998). Descriptive analysis of crashes involving pedestrians in Florida, 1990-1994. Transportation Research Record No 1636, pp. 138-145.

Bancroft G. A., Colwell D. J., Gillett J. R. (1983). A Truncated Poisson Distribution (in Notes). The Mathematical Gazette 67 (441), pp. 216-218.

Bejleri I., Steiner R.L., Kim D., 2007. GIS Methods and Tools for Bicycle and Pedestrian Crash Mapping and Analysis. In: TRB 86th Annual Meeting, Transportation Research Board, Washington DC.

Bierlaire, M., Antonini, G. and Weber, M. (2003). Behavioral dynamics for pedestrians, in K. Axhausen (ed) Moving through nets: the physical and social dimensions of travel, Elsevier.

Borgers A. and Timmermans H. J. P. (1986). City centre entry points, store location patterns and pedestrian route choice behaviour: A microlevel simulation model. Socio- Economic Planning Sciences 20 (1), pp. 25-31.

Cameron C., Triverdi P., (1998). Regression analysis of count data. Cambridge University Press, New York.

Chu, X., Guttenplan, M., Baltes, M., 2003.Whypeople cross where they do - the role of street environment. In: TRB 82nd Annual Meeting, Transportation Research Board, Washington DC.

Crane R., Crepeau R., 1998. Does neighborhood design influence travel?: A behavioral analysis of travel diary and GIS data. Transportation Research Part D 3 (4), pp. 225-238.

Daamen W., Hoogendoorn S.P., Bovy P.H.L. (2005). First-order Pedestrian Traffic Flow Theory. In the Proceedings of the 84th TRB Annual Meeting, Transportation Research Board, Washington.

Das S., Manski C.F., Manuszak M., 2005. Walk or Wait? An Empirical Analysis of Street Crossing Decisions. Journal of Applied Econometrics 20 (4), pp. 445-577.

Duncan, C., Khattak, A., Hughes, R., 2002. Effectiveness of pedestrian safety treatments for hit-along-roadway crashes. TRB 81st Annual Meeting, Transportation Research Board, Washington DC, 2002.

Fitzpatrick, K., Ullman, B., Trout, N., 2004. On-Street Pedestrian Surveys of Pedestrian Crossing Treatments. In: TRB 83rd Annual Meeting, Transportation Research Board, Washington DC.

Hamed M.M., 2001. Analysis of pedestrians' behaviour at pedestrian crossings. Safety Science 38, pp. 63-82.

Hine J., 1996. Pedestrian travel experiences: Assessing the impact of traffic on behaviour and perceptions of safety using an in-depth interview technique. Journal of Transport Geography 4 (3), pp. 179-199.

Hoogendoorn S.P., Bovy P.H.L. (2004). Pedestrian route-choice and activity scheduling theory and models. Transportation Research Part B 38, pp. 169–190.

Howarth, C.I. (1982). The need for regular monitoring of the exposure of pedestrians and cyclists to traffic. Accident Analysis and Prevention 14 (5).

Jonah B.A., Engel G.R. (1983). Measuring the relative risk of pedestrian accidents. Accident Analysis & Prevention 15, pp. 193-206.

Keall, M. (1995). Pedestrian exposure to risk of road accident in New Zealand. Accident Anaysis and Prevention 27 (5).

Kitazawa, K., Batty, M. (2004). Pedestrian Behaviour Modelling: An Application to Retail Movements using a Genetic Algorithm. 7th International Conference on Design and Decision Support Systems in Architecture and Urban Planning.

Kukla R., Kerridge J., Willis A., Hine J. (2001). PEDFLOW: Development of an Autonomous Agent Model of Pedestrian Flow. Transportation Research Record No 1774, pp. 11-17.

Lassarre S., Papadimitriou E., Yannis G., Golias J., 2007. Measuring accident risk exposure for pedestrians in different micro-environments. Accident Analysis & Prevention 39 (6), pp. 1226-1238.

Lassarre, S., Banos, A., Bodin, F., Bonnet, E., Papadimitriou, E., Zeitouni, K. 2005. The Lille case study. Deliverable 24 of the HEARTS project, INRETS.

Lassarre S., Bonnet E., Papadimitriou E., 2007. Risk exposure assessment of pedestrians in urban area using a GIS. Electronics Proceedings, http://www.ictct.org, ICTCT Extra workshop 'Road user behaviour with a special focus on vulnerable road users', Beijing, China, 2 - 3 April 2007, 16 p.

Lee C., Abdel-Aty M. (2005). Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Accident Analysis and Prevention 37, pp. 775-786.

Lee J.Y.S, Lam W.H.K. (2008). Simulating pedestrian movements at signalized crosswalks in Hong Kong. Transportation Research Part A 42, pp. 1314-1325.

Li, Y., Tsukaguchi, H., 2005. Relationships between network topology and pedestrian route choice behaviour. Journal of the Eastern Asia Society for Transportation Studies 6, pp. 241 - 248.

Miller J.S., 2000. Geographic Information Systems: Unique analytic capabilities for the traffic safety community. Transportation Research Record No. 1734, pp. 21-28.

Papadimitriou E., Yannis G., Golias J. (2009). A critical assessment of pedestrian behaviour models. Transportation Research Part F 12 (3), pp. 242-255.

Papadimitriou E., Yannis G., Golias J. (2010). A theoretical framework for modeling pedestrians crossing behaviour along a trip. Journal of Transportation Engineering 136 (10), pp. 914-924.

Parks J.R., Schofer J.L., 2006. Characterizing neighborhood pedestrian environments with secondary data. Transportation Research Part D 11, pp. 250-263.

Pulugurtha S.S., Krishnakumar V.K., Nambisan S.S, 2007. New methods to identify and rank high pedestrian crash zones: An illustration. Accident Analysis and Prevention 39, pp. 800–811.

Schlossberg M., Brown, N., 2004. Comparing transit-oriented development sites by walkability indicators. Transportation Research Record No. 1887, pp. 34-42.

Sisiopiku V. P., Akin D., 2003. Pedestrian behaviors at and perceptions towards various pedestrian facilities: an examination based on observation and survey data. Transportation Research Part F: Traffic Psychology and Behaviour 6 (4), pp. 249-274.

WHO-World Health Organization, 2006. Health effects and risks of transport systems: the HEARTS project. WHO Regional Office for Europe.

Xiaoping Z., Tingkuan Z., Mengting L. (2009). Modeling crowd evacuation of a building based on seven methodological approaches. Building and Environment 44, pp. 437-445.

Yang J., Deng W., Wang J., Li Q., Wang Z., 2006. Modeling pedestrians road crossing behavior in traffic system micro-simulation in China. Transportation Research Part A: Policy and Practice 40 (3), pp. 280-290.

**Table 1**. Description of layers and attributes of the GIS tool

| **Layer** | **Attributes** | **Notes** |
|---|---|---|
| Network characteristics | - Type of road network (pedestrian only, vehicle only, pedestrian/vehicle) | * A single column is used |
| Road characteristics | - Number of directions (integer)<br>- Road width (metres)<br>- Number of lanes per direction (integer)<br>- Presence of lane separation (1/0) | * Number of directions and number of lanes in each direction are available in the same field.<br>* Another field is available where the number of lane is estimated.<br>* Road width in metres corresponds to the maximum width of the road |
| Traffic Lights | - X/Y coordinates | * Only traffic lights on the main road network are considered |
| Pedestrian crosswalks | - Number of crosswalks per arc (integer)<br>- Type of crosswalk (zebra, pelican, bridge, tunnel)<br>- X/Y coordinates | * A separate column is created for each type of crosswalk, including the related count<br>* The coordinates are attributed to each crosswalk separately |
| Pedestrian sidewalks | - Presence of sidewalk (1/0) | * A separate column is used for each side of the road, coded to match the node-to-node directions in the road attribute table. (e.g. Dir_AB_side, Dir_BA_side for sidewalks between nodes A and B) |

**Table 2**. Frequency of number of crossings per trip

| | Number of trips | | | |
|---|---|---|---|---|
| | Side O/D 1 | Side O/D 2 | Total | |
| **Crossings per trip** | Primary | | | Secondary |
| 0 | 23 | | 23 | 0 |
| 1 | | 32 | 32 | 11 |
| 2 | 22 | | 22 | 16 |
| 3 | | 5 | 5 | 10 |
| 4 | 0 | | 0 | 8 |
| 5 | | 1 | 1 | 4 |
| 6 | 0 | | 0 | 2 |
| 7 | | 0 | 0 | 2 |
| 8 | 0 | | 0 | 1 |
| **Total trips** | 45 | 38 | 83 | 121 |
| **Total crossings** | 44 | 52 | 96 | 159 |

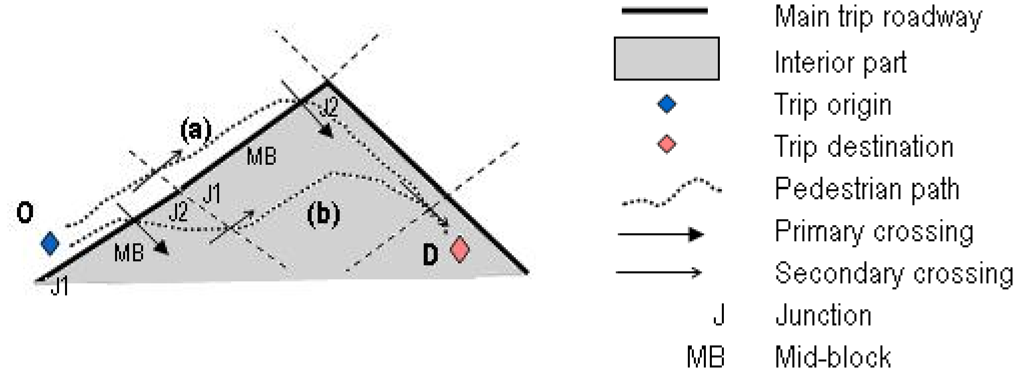


**Figure 1**. Trip characteristics, primary and secondary crossings

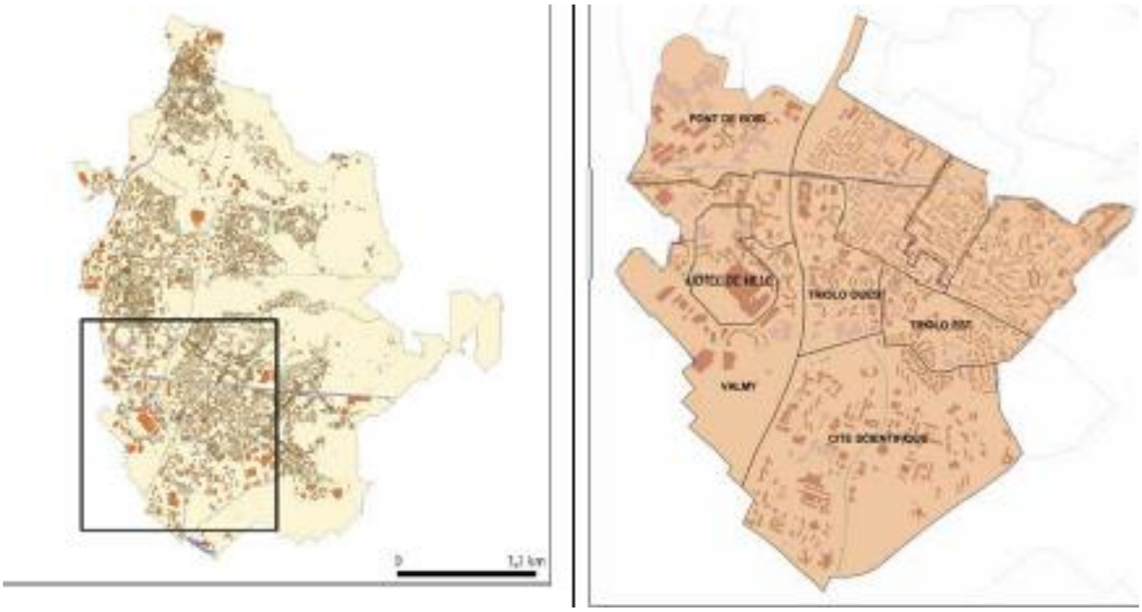


**Figure 2**. Map of Villeneuve d’Ascq - Lille and its south part.


**Figure 3**. Accessibility of the road network in the study area

**Figure 4**. Spatial matching and coding of information related to trips and crossings.

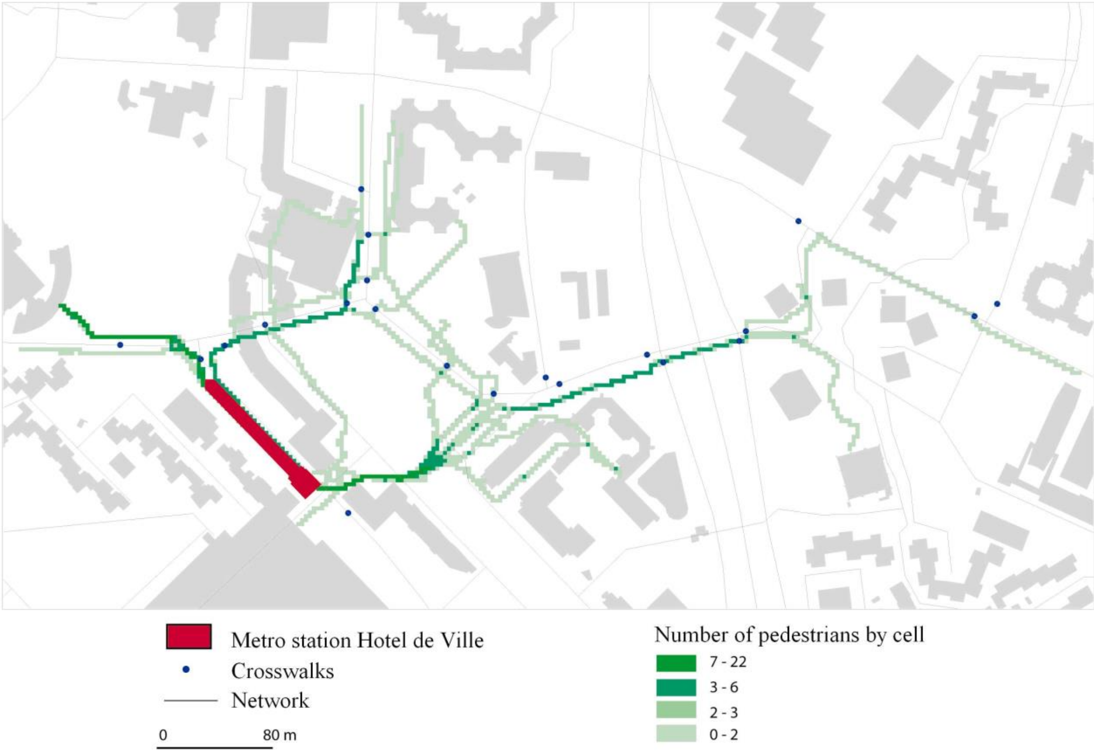


**Figure 5**. Density of sampled trips around the metro station Hotel de Ville.

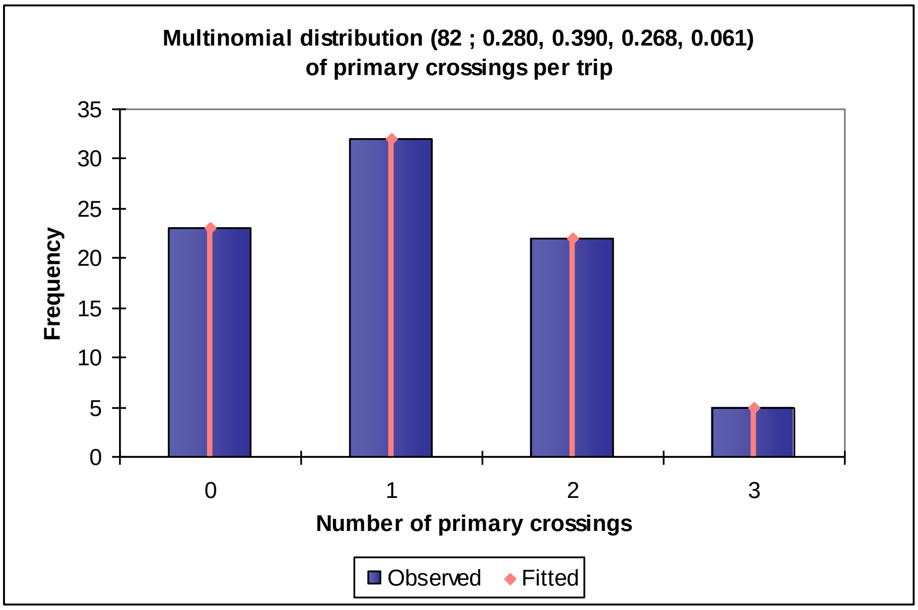


**Figure 6**. Observed and fitted distribution of the number of primary crossings per trip

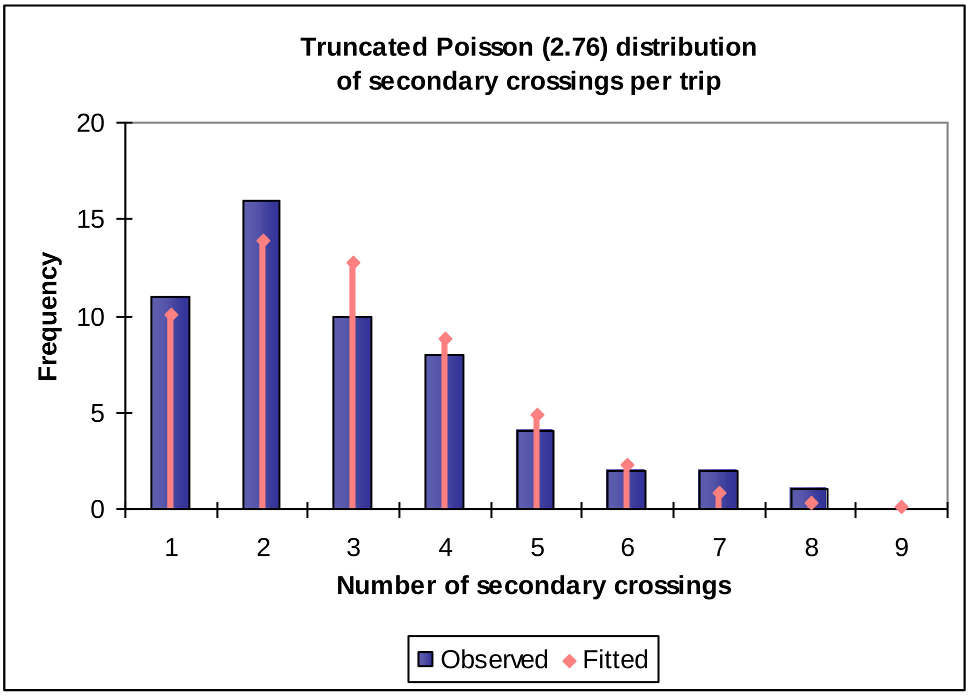


**Figure 7**. Observed and fitted distribution of the number of secondary crossings per trip

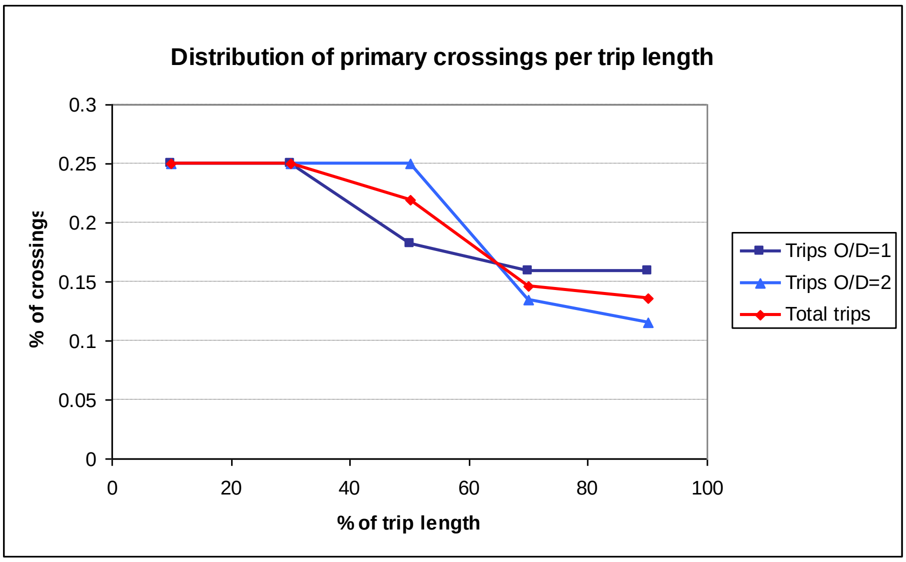


**Figure 8**. Distribution of primary crossings per trip length

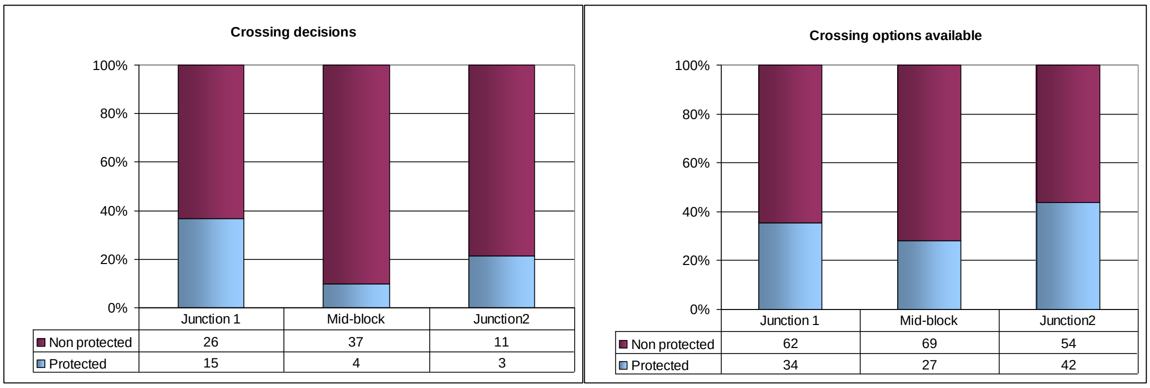


**Figure 9**. Pedestrian crossing decisions (junction or mid-block / protected or non-protected) and related options available.